

HÖLDER REGULARITY FOR PARABOLIC DE GIORGI CLASSES IN METRIC MEASURE SPACES

MATHIAS MASSON AND JUHANA SILJANDER

ABSTRACT. We give a proof for the Hölder continuity of functions in the parabolic De Giorgi classes in metric measure spaces. We assume the measure to be doubling, to support a weak $(1, p)$ -Poincaré inequality and to satisfy the annular decay property.

1. INTRODUCTION

The fine properties of the parabolic De Giorgi class are the subject of this paper. This is a class of functions which satisfy a parabolic energy estimate, which in the Euclidean case is related to the p -parabolic partial differential equation

$$-\frac{\partial u}{\partial t} + \operatorname{div}(|\nabla u|^{p-2} \nabla u) = 0,$$

and to compatible parabolic quasiminimizers. A function $u : \Omega \times (0, T) \rightarrow \mathbb{R}$, which is a parabolic K -quasiminimizer compatible with the p -parabolic partial differential equation in \mathbb{R}^d , satisfies

$$\begin{aligned} - \int_{\operatorname{supp} \phi} u \frac{\partial \phi}{\partial t} dx dt + C_1 \int_{\operatorname{supp} \phi} |\nabla u|^p dx dt \\ \leq KC_2 \int_{\operatorname{supp} \phi} |\nabla u - \nabla \phi|^p dx dt, \end{aligned}$$

with some $C_1, C_2 > 0$ and $K \geq 1$, for every smooth compactly supported function ϕ in $\Omega \times (0, T)$. Here Ω denotes a domain in \mathbb{R}^d .

Our main result is the local Hölder continuity of the parabolic De Giorgi class functions in metric measure spaces, extending the results obtained for parabolic K -quasiminima by Zhou [Zho93, Zho94] in \mathbb{R}^d with the Lebesgue measure. Historically, the parabolic version of De Giorgi classes has been investigated in euclidean spaces by Ladyzhenskaja- Solonnikov-Ural'ceva [LSU68] and DiBenedetto, and later by [DiB88], Wieser [Wie87] and Gianazza-Vespri [GV06], to name a few.

Our argument is a modification of the DiBenedetto scheme [DiB86, DiB93, Urb08]. We assume the underlying measure to be doubling and to support a weak $(1, p)$ -Poincaré inequality. Together these imply a Sobolev inequality. Also, we assume the metric measure space to satisfy the annular decay property [Buc99]. In order to study general measure spaces instead of proving

2010 *Mathematics Subject Classification.* Primary 35B65. Secondary 35K65, 31E05.

Key words and phrases. Hölder continuity, intrinsic scaling, Moser iteration, De Giorgi method, Newtonian space, upper gradient, quasiminima, De Giorgi class, nonlinear parabolic equations, analysis on metric spaces.

the argument only for the usual Lebesgue measure, we base the proof on integral averages, so that the actual scaling properties of the measure are not needed. A similar technique is applied in [KLSU11, KSU11]. In the elliptic case, the weighted theory has originally been studied for example by Chiarenza and Serapioni, [CS84a, CS84b, CS85].

Due to the general nature of the argument, our approach may open possibilities to establish Hölder continuity in some specific cases of interest, such as for solutions of sub-parabolic equations in Heisenberg groups. Indeed, as the Heisenberg group equipped with the Haar measure is a length space, it is known to satisfy the α -annular decay property [Buc99].

Our argument has several similarities with the Euclidean case, but the new material is substantial. Since standard gradients cannot be defined in a general metric space, we consider upper gradients, and replace the standard Sobolev spaces with Newtonian spaces, see [BB11, Sha00]. The parabolic De Giorgi class is defined using the upper gradients, and the modified DiBenedetto method is carried out accordingly.

The motive behind defining the De Giorgi class in this way, is that it enables us to extend the study of parabolic partial differential equations to metric measure spaces. This in turn helps us to better understand those aspects of the theory which are independent of the geometry of the space, where the partial differential equation is originally defined. In practice the extension is done via concepts which in \mathbb{R}^d are closely related to the PDE at hand, and are definable without assuming Euclidean structure of the underlying space.

For the elliptic case such a concept are the quasiminima, which are known to belong to the elliptic De Giorgi classes, see [Giu, Gia93, KS01]. In \mathbb{R}^d , classical examples of functions belonging to parabolic De Giorgi classes are the solutions of parabolic partial differential equations as well as the parabolic quasiminima. The latter has been studied by Zhou in [Zho93, Zho94] and by Wieser in [Wie87], whereas for PDEs we refer to [DiB93, Urb08].

In the last part of this paper we show that parabolic quasiminima in metric measure spaces belong to the parabolic De Giorgi class. A somewhat unexpected difficulty arises in proving the usual De Giorgi estimates for parabolic quasiminima in metric spaces. Indeed, since taking upper gradients is not a linear operation, the usual time mollification argument used in the Euclidean case seems to be destroyed. We circumvent this by introducing so called Cheeger derivatives [Che99], which have the property that taking a Cheeger derivative is a linear operation.

2. PRELIMINARIES

2.1. Metric measure spaces. Let (X, d, μ) be a complete metric measure space with metric d and a positive complete Borel measure μ . The measure μ is said to be doubling if there exists a universal constant $C_\mu \geq 1$ such that

$$\mu(B(x, 2r)) \leq C_\mu \mu(B(x, r)),$$

for every $r > 0$ and $x \in X$. Here $B(x, r)$ denotes the standard open ball

$$B(x, r) = \{y \in X : d(x, y) < r\}.$$

The dimension related to the doubling measure is defined to be

$$d_\mu = \log_2 C_\mu.$$

By iterating the doubling condition, it follows that

$$\frac{\mu(B(z, r))}{\mu(B(y, R))} \geq C_\mu^{-2} \left(\frac{r}{R} \right)^{d_\mu},$$

for all balls $B(y, R) \subset X$, $z \in B(y, R)$ and $0 < r \leq R < \infty$. Given $\alpha > 0$ and a metric space (X, d, μ) with a doubling μ , we say that the space satisfies the α -annular decay property if there exists a constant $c \geq 1$ such that

$$(2.1) \quad \mu(B(x, r) \setminus B(x, (1 - \delta)r)) \leq c\delta^\alpha \mu(B(x, r))$$

for all $B(x, r) \subset X$ and for all $0 < \delta < 1$. Every length space has this property, in particular this is true in \mathbb{R}^d . For further information about the spaces which satisfy this, see [Buc99].

2.2. Upper gradients. Let $\Omega \subset X$ be open. Following [HK98], a non-negative Borel measurable function $g : \Omega \rightarrow [0, \infty]$ is said to be an upper gradient of a function $u : \Omega \rightarrow [-\infty, \infty]$ in Ω , if for all compact rectifiable paths γ joining x and y in Ω we have

$$(2.2) \quad |u(x) - u(y)| \leq \int_\gamma g \, ds.$$

In case $u(x) = u(y) = \infty$ or $u(x) = u(y) = -\infty$, the left side is defined to be ∞ . Assume $1 \leq p < \infty$. The p -modulus of a family of paths Γ in Ω is defined to be

$$\inf_\rho \int_\Omega \rho^p \, d\mu,$$

where the infimum is taken over all non-negative Borel measurable functions ρ such that for all rectifiable paths γ which belong to Γ , we have

$$\int_\gamma \rho \, ds \geq 1.$$

A property is said to hold for p -almost all paths, if the set of non-constant paths for which the property fails is of zero p -modulus. Following [KM98, Sha00], if (2.2) holds for p -almost all paths γ in X , then g is said to be a p -weak upper gradient of u .

When $1 < p < \infty$ and $u \in L^p(\Omega)$, it can be shown [Sha01] that there exists a minimal p -weak upper gradient of u , we denote it by g_u , in the sense that g_u is a p -weak upper gradient of u and for every p -weak upper gradient g of u it holds $g_u \leq g$ μ -almost everywhere in Ω . Moreover, if $v = u$ μ -almost everywhere in a Borel set $A \subset \Omega$, then $g_v = g_u$ μ -almost everywhere in A . Also, if $u, v \in L^p(\Omega)$, then μ -almost everywhere in Ω , we have

$$\begin{aligned} g_{u+v} &\leq g_u + g_v, \\ g_{uv} &\leq |u|g_v + |v|g_u. \end{aligned}$$

Proofs for these properties and more on upper gradients in metric spaces can be found for example in [BB11] and the references therein.

2.3. Newtonian spaces. Following [Sha00], for $1 < p < \infty$ and $u \in L^p(\Omega)$, we define

$$\|u\|_{1,p,\Omega} = \|u\|_{L^p(\Omega,\mu)} + \|g_u\|_{L^p(\Omega,\mu)},$$

and

$$\tilde{N}^{1,p}(\Omega) = \{u : \|u\|_{1,p,\Omega} < \infty\}.$$

An equivalence relation in $\tilde{N}^{1,p}(\Omega)$ is defined by saying that $u \sim v$ if

$$\|u - v\|_{\tilde{N}^{1,p}(\Omega)} = 0.$$

The *Newtonian space* $N^{1,p}(\Omega)$ is defined to be the space $\tilde{N}^{1,p}(\Omega)/\sim$, with the norm

$$\|u\|_{N^{1,p}(\Omega)} = \|u\|_{1,p,\Omega}.$$

A function u belongs to the local Newtonian space $N_{\text{loc}}^{1,p}(\Omega)$ if it belongs to $N^{1,p}(\Omega')$ for every $\Omega' \subset\subset \Omega$. For more properties of Newtonian spaces, see [Hei01, Sha00, BB11].

2.4. Poincaré's inequality. For positive $1 \leq q < \infty$, $1 \leq p < \infty$, the measure μ is said to support a weak (q, p) -Poincaré inequality if there exist constants $P_0 > 0$ and $\tau \geq 1$ such that

$$(2.3) \quad \left(\int_{B(x,r)} |v - v_{B(x,r)}|^q d\mu \right)^{1/q} \leq P_0 r \left(\int_{B(x,\tau r)} g_v^p d\mu \right)^{1/p},$$

for every $v \in N^{1,p}(X)$ and $B(x, \tau r) \subset X$. Here we use the notation

$$v_{B(x,r)} = \int_{B(x,r)} v d\mu = \frac{1}{\mu(B(x,r))} \int_{B(x,r)} v d\mu.$$

In case $\tau = 1$, we say a (q, p) -Poincaré inequality is in force. In a general metric measure space setting, it is of interest to have assumptions which are invariant under bi-Lipschitz mappings. The weak (q, p) -Poincaré inequality has this quality.

For a metric space X equipped with a doubling measure μ , the following result of [HK95] is known: If X supports a weak $(1, p)$ -Poincaré inequality for some $1 < p < \infty$, then X also supports a weak (κ, p) -Poincaré inequality, where

$$\kappa = \begin{cases} \frac{d_\mu p}{d_\mu - p}, & \text{for } 1 < p < d_\mu, \\ 2p, & \text{otherwise,} \end{cases}$$

possibly with different constants $P'_0 > 0$ and $\tau' \geq 1$. As a consequence of this, by the (κ, p) -Poincaré inequality and for example by Proposition 5.41 in [BB11], there exists a positive constant C such that

$$(2.4) \quad \left(\int_{B(x,r)} |v|^\kappa d\mu \right)^{1/\kappa} \leq C r \left(\int_{B(x,r)} g_v^p d\mu \right)^{1/p}.$$

for every $v \in N_0^{1,p}(X)$ and $B(x, r) \subset X$.

Remark 2.5. It is a result of [KZ08], that when $1 < p < \infty$ and (X, d) is a complete metric space with doubling measure μ , the weak $(1, p)$ -Poincaré inequality implies a weak $(1, q)$ -Poincaré inequality for some $1 < q < p$. Then by the above discussion, X also supports a weak (κ, q) -Poincaré inequality with some $\kappa > q$. By Hölder's inequality, the left hand side of the weak (κ, q) -Poincaré inequality can be estimated from below by replacing κ with any positive $\kappa' < \kappa$. Hence we conclude, that if X supports a weak $(1, p)$ -Poincaré inequality with $1 < p < \infty$, then X also supports a weak (q, q) -Poincaré inequality with some $1 < q < p$.

2.5. Parabolic upper gradients and Newtonian spaces. We define the parabolic Newtonian space $L^p(0, T; N^{1,p}(\Omega))$ to be the space of functions $u(x, t)$ such that for almost every $0 < t < T$ the function $u(\cdot, t)$ belongs to $N^{1,p}(\Omega)$, and

$$\int_0^T \|u(\cdot, t)\|_{1,p,\Omega}^p dt < \infty.$$

The definition of the space $L_{\text{loc}}^p(0, T; N_{\text{loc}}^{1,p}(\Omega))$ is obvious. In what follows we will denote the product measure by $d\nu = d\mu dt$. Let $u \in L_{\text{loc}}^p(0, T; N_{\text{loc}}^{1,p}(\Omega))$. The parabolic minimal p -weak upper gradient of u is defined in a natural way by setting

$$g_u(x, t) = g_{u(\cdot, t)}(x),$$

at ν -almost every $(x, t) \in \Omega \times (0, T) = \Omega_T$. For the sake of conciseness we refer to the parabolic minimal p -weak upper gradient of a time dependent function, by calling it the upper gradient.

Next we define the class of functions for which we prove local Hölder continuity.

Definition 2.6 (Parabolic De Giorgi class). For $p > 2$, we say that $u \in L_{\text{loc}}^p(0, T; N_{\text{loc}}^{1,p}(\Omega))$ belongs to the *parabolic De Giorgi class*, if there exists a positive constant C , with which we have

$$\begin{aligned} & \text{ess sup}_{\tau_1 < t < \tau_0} \int_{B(x, r_1)} (u(x, t) - k)_{\pm}^2 d\mu + \int_{\tau_1}^{\tau_0} \int_{B(x, r_1)} g_{(u-k)_{\pm}}^p d\nu \\ & \leq \frac{C}{(r_2 - r_1)^p} \int_{\tau_2}^{\tau_0} \int_{B(x, r_2)} (u - k)_{\pm}^p d\nu + \frac{C}{(\tau_1 - \tau_2)} \int_{\tau_2}^{\tau_1} \int_{B(x, r_2)} (u - k)_{\pm}^2 d\nu, \end{aligned}$$

for every $x \in \Omega$, $r_1 < r_2$ and $\tau_2 < \tau_1 < \tau_0$, such that $B(x, r_2) \times (\tau_2, \tau_0) \subset \Omega_T$. As an immediate consequence, by the Lebesgue differentiation theorem, we see by taking the limit $\tau_2 \rightarrow \tau_1$, that for almost every $\tau_1 \in (0, \tau_0)$, we have

$$\begin{aligned} (2.7) \quad & \text{ess sup}_{\tau_1 < t < \tau_0} \int_{B(x, r_1)} (u(x, t) - k)_{\pm}^2 d\mu + \int_{\tau_1}^{\tau_0} \int_{B(x, r_1)} g_{(u-k)_{\pm}}^p d\nu \\ & \leq \frac{C}{(r_2 - r_1)^p} \int_{\tau_1}^{\tau_0} \int_{B(x, r_2)} (u - k)_{\pm}^p d\nu + C \int_{B(x, r_2)} (u(x, \tau_1) - k)_{\pm}^2 d\mu. \end{aligned}$$

The following regularity theorem is the main result of this paper.

Theorem 2.8. Assume (X, d) is a complete metric space equipped with a complete doubling positive Borel measure μ . Assume X supports a weak

$(1, p)$ -Poincaré inequality for $2 < p < \infty$, and satisfies the α -annular decay property for some $0 < \alpha < 1$. Let $u \in L_{loc}^p(0, T; N_{loc}^{1,p}(\Omega))$ be locally essentially bounded and belong to the parabolic De Giorgi class. Then u has a representative which is locally Hölder continuous in Ω_T .

The proof is based on the DiBenedetto argument and it is divided into two alternatives. We will use intrinsic scaling [Urb08] in order to homogenize the powers on the right hand side of the estimate in definition 2.6. We start by constructing the geometry.

2.6. Initial setting. Let $\lambda \geq 1$ be an a priori constant, which has been fixed in a manner which will become clear later in this text. Assume a compact set $S \subset \Omega_T$. Our aim is to show that u is Hölder continuous in S . Since S is compact, there exists an open set F such that $S \subset F \subset \subset \Omega_T$.

Since u is assumed to be locally essentially bounded, $\text{ess osc}_F u$ is finite. We assume that $\text{ess osc}_F u > 0$, because otherwise Hölder continuity in S is trivially true. We redefine $u(x, t) = \text{ess sup}_F u$ in the ν -negligible subset of F where $u(x, t) \notin [\text{ess inf}_F u, \text{ess sup}_F u]$. Set

$$\theta_{\pm} = \left(\frac{\text{ess osc}_F u}{\gamma_{\pm}} \right)^{2-p},$$

where $\gamma_- = 2$ and $\gamma_+ = 2^\lambda$. Let r be a positive number such that for every $(x, t) \in S$ we have $B(x, 2\tau r) \times (t - 2\theta_+ r^p, t) \subset F$. Clearly r is controlled by the distance of S to the complement of Ω_T .

Considering any point $(x, t) \in S$, it will turn out that the constants in the reduction of oscillation of u in the neighborhood of (x, t) depend only on $p, d_\mu, C_\mu, P_0, \tau$ and on the constant r . This in turn implies by a standard iterative argument [Urb08], that u is Hölder continuous in S , and that the modulus of Hölder continuity will depend only on $p, d_\mu, C_\mu, P_0, \tau, r$ and on $\text{ess osc}_F u$. Therefore, to prove the Hölder continuity of u in a compact set S , it is enough to consider any one point in S , and examine the oscillation of u in its neighborhood.

So from here on, let (x_0, t_0) denote some fixed point in S . For brevity, in what follows we will refer to the set of constants $p, d_\mu, C_\mu, P_0, \tau$ as the data.

3. ESTIMATES FOR THE PARABOLIC DE GIORGI CLASS

Let $t^* \in (t_0 - \theta_+ r^p, t_0)$. For $n = 0, 1, \dots$ we denote

$$r_n = \frac{r_0}{2} + \frac{r_0}{2^{n+1}}, \quad Q_n^\pm = B_n \times T_n^\pm = B(x_0, r_n) \times (t^* - \theta_\pm r_n^p, t^*),$$

and

$$A_n^\pm = \{(x, t) \in Q_n^\pm : (u(x, t) - k_n^\pm)_\pm > 0\},$$

where r_0 and k_n^\pm will be chosen later, according to the situation at hand. The following two lemmas are of central importance to the proof.

Lemma 3.1. *Let $(k_n^+)_n$ be an increasing sequence and $(k_n^-)_n$ a decreasing sequence of real numbers. Let u belong to the parabolic De Giorgi class and assume that for some number $0 < \varepsilon < 1$, we have*

$$(3.2) \quad (u - k_n^\pm)_\pm \leq \varepsilon \text{ess osc}_F u \quad \text{and} \quad |k_{n+1}^\pm - k_n^\pm| \geq \frac{\varepsilon \text{ess osc}_F u}{2^{n+2}},$$

Then there exists a constant C_0 which depends only on the data such that

$$\frac{\nu(A_{n+2}^\pm)}{\nu(Q_{n+2}^\pm)} \leq C_0(3 + (\varepsilon\gamma_\pm)^{p-2} + (\varepsilon\gamma_\pm)^{2-p})^{2-p/\kappa} 4^{pn(2-p/\kappa)} \left(\frac{\nu(A_n^\pm)}{\nu(Q_n^\pm)} \right)^{2-p/\kappa}$$

for every $n = 0, 1, \dots$.

Proof. For each n , let $\varphi_n \in \text{Lip}(\Omega)$, $0 \leq \varphi_n \leq 1$, be a cut off function such that $\varphi_n(x) = 1$ in B_{n+1} , the support of φ_n is a compact subset of B_n , and $g_{\varphi_n} \leq 2^{n+2}/r_0$. Let κ be the constant from (2.4). By Hölder's inequality we can write

$$\begin{aligned} \int_{Q_{n+2}^\pm} (u - k_n^\pm)_\pm^{p(2-p/\kappa)} d\nu &\leq \int_{T_{n+2}^\pm} \left(\int_{B_{n+2}} (u - k_n^\pm)_\pm^p d\mu \right)^{(\kappa-p)/\kappa} \\ &\quad \times \left(\int_{B_{n+2}} (u - k_n^\pm)_\pm^\kappa d\mu \right)^{p/\kappa} dt \\ &\leq \frac{|T_{n+1}^\pm| \mu(B_{n+1})}{|T_{n+2}^\pm| \mu(B_{n+2})} \left(\text{ess sup}_{T_{n+1}^\pm} \int_{B_{n+1}} (u - k_n^\pm)_\pm^p d\mu \right)^{(\kappa-p)/\kappa} \\ &\quad \times \int_{T_{n+1}^\pm} \left(\int_{B_{n+1}} ((u - k_n^\pm)_\pm \varphi_{n+1})^\kappa d\mu \right)^{p/\kappa} dt. \end{aligned}$$

By the doubling property of μ , the measure factor on the right hand side is uniformly bounded for every n . Now, since $(u - k_n^\pm)_\pm \varphi_{n+1} \in N_0^{1,p}(B_{n+1})$, we can use inequality (2.4) to obtain

$$\begin{aligned} \int_{T_{n+1}^\pm} \left(\int_{B_{n+1}} ((u - k_n^\pm)_\pm \varphi_{n+1})^\kappa d\mu \right)^{p/\kappa} dt &\leq Cr_0^p \int_{Q_{n+1}^\pm} g_{(u-k_n^\pm)_\pm \varphi_{n+1}}^p d\nu \\ &\leq Cr_0^p \int_{Q_{n+1}^\pm} g_{(u-k_n^\pm)_\pm}^p d\nu + C2^{p(n+2)} \int_{Q_{n+1}^\pm} (u - k_n^\pm)_\pm^p d\nu \\ &\leq Cr_0^p \int_{Q_{n+1}^\pm} g_{(u-k_n^\pm)_\pm}^p d\nu + C2^{p(n+2)} (\varepsilon \text{ess osc}_F u)^p \frac{\nu(A_n^\pm)}{\nu(Q_n^\pm)}. \end{aligned}$$

Next we use the definition of the parabolic De Giorgi class to estimate the first integral term on the right hand side of the above expression. We have

$$\begin{aligned} r_0^p \int_{Q_{n+1}^\pm} g_{(u-k_n^\pm)_\pm}^p d\nu &\leq C \left(2^{pn} \int_{Q_n^\pm} (u - k_n^\pm)_\pm^p d\nu + \frac{2^{pn}}{\theta_\pm} \int_{Q_n^\pm} (u - k_n^\pm)_\pm^2 d\nu \right) \\ &\leq C2^{pn} \left(1 + \frac{1}{\theta_\pm} (\varepsilon \text{ess osc}_F u)^{2-p} \right) (\varepsilon \text{ess osc}_F u)^p \frac{\nu(A_n^\pm)}{\nu(Q_n^\pm)} \\ &= C2^{pn} \left(1 + (\gamma_\pm \varepsilon)^{2-p} \right) (\varepsilon \text{ess osc}_F u)^p \frac{\nu(A_n^\pm)}{\nu(Q_n^\pm)}. \end{aligned}$$

Similarly, since u belongs to the parabolic De Giorgi class, by the doubling property of μ and by the fact that $\mu(B_n)\theta_\pm r_n^p = \nu(Q_n^\pm)$, we have

$$\begin{aligned} \operatorname{ess\,sup}_{T_{n+1}^\pm} \int_{B_{n+1}} (u - k_n^\pm)_\pm^p d\mu &\leq \frac{(\gamma_\pm \varepsilon)^{p-2}}{\theta_\pm} \operatorname{ess\,sup}_{T_{n+1}^\pm} \int_{B_{n+1}} (u - k_n^\pm)_\pm^2 d\mu \\ &\leq C(\gamma_\pm \varepsilon)^{p-2} \left(2^{pn} \int_{Q_n^\pm} (u - k_n^\pm)_\pm^p d\nu + \frac{2^{pn}}{\theta_\pm} \int_{Q_n^\pm} (u - k_n^\pm)_\pm^2 d\nu \right) \\ &\leq C2^{pn} ((\gamma_\pm \varepsilon)^{p-2} + 1) (\varepsilon \operatorname{ess\,osc}_F u)^p \frac{\nu(A_n^\pm)}{\nu(Q_n^\pm)}. \end{aligned}$$

Collecting the obtained estimates yields,

$$\begin{aligned} \int_{Q_{n+2}^\pm} (u - k_n^\pm)_\pm^{p(2-p/\kappa)} d\nu \\ \leq \left(C2^{pn} (3 + (\varepsilon\gamma_\pm)^{p-2} + (\varepsilon\gamma_\pm)^{2-p}) (\varepsilon \operatorname{ess\,osc}_F u)^p \frac{\nu(A_n^\pm)}{\nu(Q_n^\pm)} \right)^{2-p/\kappa}. \end{aligned}$$

On the other hand, by (3.2) and since μ is doubling, we have

$$\begin{aligned} \int_{Q_{n+2}^\pm} (u - k_n^\pm)_\pm^{p(2-p/\kappa)} d\nu &\geq |k_{n+1}^\pm - k_n^\pm|^{p(2-p/\kappa)} \frac{\nu(A_{n+2}^\pm)}{\nu(Q_{n+2}^\pm)} \\ &\geq \left(\frac{\varepsilon \operatorname{ess\,osc}_F u}{2^{n+2}} \right)^{p(2-p/\kappa)} \frac{\nu(A_{n+2}^\pm)}{\nu(Q_{n+2}^\pm)}. \end{aligned}$$

Therefore, for each $n \geq 0$ we have

$$\frac{\nu(A_{n+2}^\pm)}{\nu(Q_{n+2}^\pm)} \leq C4^{pn(2-p/\kappa)} (3 + (\varepsilon\gamma_\pm)^{p-2} + (\varepsilon\gamma_\pm)^{2-p})^{2-p/\kappa} \left(\frac{\nu(A_n^\pm)}{\nu(Q_n^\pm)} \right)^{2-p/\kappa},$$

where C depends only on the data. \square

We also prove the following time independent variant of the above lemma.

Lemma 3.3. *Suppose the assumptions of Lemma 3.1 hold. Suppose in addition that for some $t_0 - 2\theta_+ r_0^p < t' < t_0$, where t' is a Lebesgue point of the mapping*

$$t \mapsto \int_{B(x_0, r)} (u(x, t) - k)_\pm^2 d\mu,$$

we have $(u - k_n^\pm)_\pm(x, t') = 0$ at μ -almost every $x \in B(x_0, r_n)$ for every $n \geq 0$. Then for the time independent sequence

$$Q_n^\pm = B(x_0, r_n) \times (t', t_0),$$

and corresponding sets A_n^\pm , there exists a constant C_0 , which depends only on the data, such that

$$\frac{\nu(A_{n+2}^\pm)}{\nu(Q_{n+2}^\pm)} \leq C_0 \left(2 + 2^{(\lambda+1)(p-2)} \right)^{2-p/\kappa} 4^{pn(2-p/\kappa)} \left(\frac{\nu(A_n^\pm)}{\nu(Q_n^\pm)} \right)^{2-p/\kappa}$$

for every $n = 0, 1, 2, \dots$.

Proof. We proceed exactly as in the proof of Lemma 3.1, with $T_n^\pm = (t', t_0)$ for every n , but now instead of the parabolic De Giorgi class, we can use (2.7) and the assumption that we have $(u - k_n^\pm)_\pm(x, t') = 0$ for μ -almost every $x \in B(x_0, r_0)$. We estimate

$$r_0^p \int_{Q_{n+1}^\pm} g_{(u-k_n^\pm)_\pm}^p d\nu \leq C 2^{pn} \int_{Q_n^\pm} (u - k_n^\pm)_\pm^p d\nu \leq C 2^{pn} (\varepsilon \operatorname{ess\,osc}_F u)^p \frac{\nu(A_n^\pm)}{\nu(Q_n^\pm)},$$

and

$$\begin{aligned} \operatorname{ess\,sup}_{T_{n+1}^\pm} \int_{B_{n+1}} (u - k_n^\pm)_\pm^p d\mu &\leq C (\varepsilon \operatorname{ess\,osc}_F u)^{p-2} \frac{2^{pn}(t_0 - t')}{r_0^p} \int_{Q_n^\pm} (u - k_n^\pm)_\pm^p d\nu \\ &\leq C 2^{pn} (\varepsilon 2\gamma_+)^{p-2} (\varepsilon \operatorname{ess\,osc}_F u)^p \frac{\nu(A_n^\pm)}{\nu(Q_n^\pm)}. \end{aligned}$$

Plugging these into the proof of Lemma 3.1 yields

$$\begin{aligned} &\int_{Q_{n+2}^\pm} (u - k_n^\pm)_\pm^{p(2-p/\kappa)} d\nu \\ &\leq \left(C 2^{pn} (2 + (\varepsilon 2\gamma_+)^{p-2}) (\varepsilon \operatorname{ess\,osc}_F u)^p \frac{\nu(A_n^\pm)}{\nu(Q_n^\pm)} \right)^{2-p/\kappa}. \end{aligned}$$

As in the proof of Lemma 3.1, after noting that $\varepsilon < 1$, this leads to the estimate

$$\frac{\nu(A_{n+2}^\pm)}{\nu(Q_{n+2}^\pm)} \leq C 4^{pn(2-p/\kappa)} \left(2 + 2^{(\lambda+1)(p-2)} \right)^{2-p/\kappa} \left(\frac{\nu(A_n^\pm)}{\nu(Q_n^\pm)} \right)^{2-p/\kappa}.$$

□

Later in this proof, depending on the situation, we will use Lemma 3.1 or Lemma 3.3 together with the following real analytic lemma.

Lemma 3.4. *Let $(Y_n)_n$ be a sequence of positive numbers, satisfying*

$$Y_{n+1} \leq C b^n Y_n^{1+\alpha}$$

where $C, b > 1$ and $\alpha > 0$. Then Y_n converges to 0 as $n \rightarrow \infty$, provided that

$$(3.5) \quad Y_0 \leq C^{-1/\alpha} b^{1-\alpha^2}.$$

Proof. For the proof, we refer to [DiB93].

□

By Lemma 3.4, once the requirements of Lemma 3.1 or Lemma 3.3 have been established, the convergence to zero of $\nu(A_{2n}^\pm)$ follows, provided we can first show that the corresponding initial condition (3.5) is satisfied.

Next we divide the proof in two complementary alternatives and study them separately.

4. THE FIRST ALTERNATIVE

Recall that $\gamma_- = 2$, and set

$$(4.1) \quad k_n^- = \operatorname{ess\,inf}_F u + \frac{\operatorname{ess\,osc}_F u}{4} + \frac{\operatorname{ess\,osc}_F u}{2^{n+2}} \quad \text{and} \quad r_0 = r.$$

Then

$$(u - k_n^-)_- \leq \frac{\operatorname{ess\,osc}_F u}{2} \quad \text{and} \quad |k_{n+1}^\pm - k_n^\pm| \geq \frac{1}{2} \frac{\operatorname{ess\,osc}_F u}{2^{n+2}},$$

and so we see that here $\varepsilon = 1/2$. Plug these into Lemma 3.1, and define

$$\alpha_0 = (C_0(3 + 1^{p-2} + 1^{2-p})^{2-p/\kappa})^{-1/(1-p/\kappa)} (4^{p2(2-p/\kappa)})^{1-(1-p/\kappa)^2}.$$

Note that α_0 does not depend on λ .

Suppose there exists a $t^* \in (t_0 - \theta_+ r^p, t_0)$ such that

$$(4.2) \quad \begin{aligned} & \nu(\{B(x_0, r) \times (t^* - \theta_- r^p, t^*) : u(x, t) \leq k_0^-\}) \\ & < \alpha_0 \nu(B(x_0, r) \times (t^* - \theta_- r^p, t^*)). \end{aligned}$$

We refer to this condition as *the first alternative*. Using the abbreviations introduced in the previous section we can write inequality (4.2) as

$$\frac{\nu(A_0^-)}{\nu(Q_0^-)} < \alpha_0.$$

By Lemma 3.1 and Lemma 3.4 this implies that $\nu(A_{2n}^-)/\nu(Q_{2n}^-) \rightarrow 0$ as $n \rightarrow 0$. Hence if the first alternative is in force, we obtain

$$u(x, t) \geq \operatorname{ess\,inf}_F u + \frac{\operatorname{ess\,osc}_F u}{4}$$

ν -a.e. in $B(x_0, \frac{r}{2}) \times (t^* - \theta_- (\frac{r}{2})^p, t^*)$. This implies that we can fix a time level $t' \in (t_0 - 2\theta_+ r^p, t_0 - \theta_- (\frac{r}{4})^p)$ at which we have

$$(4.3) \quad u(x, t') \geq \operatorname{ess\,inf}_F u + \frac{\operatorname{ess\,osc}_F u}{4} \text{ for } \mu\text{-a.e. } x \in B(x_0, r/2).$$

Now we are able to use (2.7) to obtain the following.

Lemma 4.4. *Assume the first alternative. Then there exists a constant C , which depends only on the data, such that for all $s > 1$ we have*

$$\frac{\nu(\{B(x_0, r/4) \times (t', t_0) : u < \operatorname{ess\,inf}_F u + \frac{\operatorname{ess\,osc}_F u}{2^{\lambda+s}}\})}{\nu(B(x_0, r/4) \times (t', t_0))} \leq \frac{C}{2^{s(p-2)}}.$$

Proof. We set

$$k = \operatorname{ess\,inf}_F u + \frac{\operatorname{ess\,osc}_F u}{2^{\lambda+s-1}},$$

where $s > 1$. By (4.3) we have $(u - k)_-(x, t') = 0$ for μ -a.e. $x \in B(x_0, r/2)$, and so by (2.7), we obtain

$$(4.5) \quad \operatorname{ess\,sup}_{t' < t < t_0} \int_{B(x_0, \frac{r}{4})} (u - k)_-^2(x, t) d\mu \leq \frac{C}{r^p} \int_{t'}^{t_0} \int_{B(x_0, \frac{r}{2})} (u - k)_-^p d\mu dt.$$

On the other hand, for each $t \in (t', t_0)$ in the set

$$\left\{ x \in B(x_0, r/4) : u(x, t) < \operatorname{ess\,inf}_F u + \frac{\operatorname{ess\,osc}_F u}{2^{\lambda+s}} \right\},$$

we have

$$(u - k)_-(x, t) \geq \frac{\operatorname{ess\,osc}_F u}{2^{\lambda+s}}.$$

We use this with inequality (4.5) to conclude that for almost every $t \in (t', t_0)$

$$\begin{aligned} & \left(\frac{\operatorname{ess\,osc}_F u}{2^{\lambda+s}} \right)^2 \mu \left(\left\{ B(x_0, r/4) : u(\cdot, t) < \operatorname{ess\,inf}_F u + \frac{\operatorname{ess\,osc}_F u}{2^{\lambda+s}} \right\} \right) \\ & \leq \frac{C}{r^p} \left(\frac{\operatorname{ess\,osc}_F u}{2^{\lambda+s-1}} \right)^p |t_0 - t'| \mu(B(x_0, r/4)). \end{aligned}$$

Since

$$|t_0 - t'| < \left(\frac{\text{ess osc}_F u}{2^\lambda} \right)^{2-p} 2r^p,$$

we can further estimate the right hand side to obtain

$$\mu \left(\left\{ B(x_0, r/4) : u(x, t) < \text{ess inf}_F u + \frac{\text{ess osc}_F u}{2^{\lambda+s}} \right\} \right) \leq \frac{C \mu(B(x_0, r/4))}{2^{s(p-2)}},$$

for almost all $t \in (t', t_0)$, where C depends only on the constant in Lemma 3.3. This implies

$$\frac{\nu(\{B(x_0, r/4) \times (t', t_0) : u < \text{ess inf}_F u + \frac{\text{ess osc}_F u}{2^{\lambda+s}}\})}{\nu(B(x_0, r/4) \times (t', t_0))} \leq \frac{C}{2^{s(p-2)}},$$

as desired. \square

We can now use the above Lemma together with Lemma 3.3 to guarantee that the initial conditions needed to employ Lemma 3.4 are met in a cylinder with upper time level at t_0 .

Lemma 4.6. *Assume the first alternative. Then there exists a positive integer s which depends only on the data and on λ , such that*

$$u \geq \text{ess inf}_F u + \frac{\text{ess osc}_F u}{2^{s+1}} \quad \text{a.e. in } B(x_0, r/8) \times (t_0 - \theta_-(r/8)^p, t_0).$$

Proof. In Lemma 3.3, set

$$r_0 = \frac{r}{8}, \quad r_n = \frac{r}{8} + \frac{r}{8 \cdot 2^n},$$

$$Q_n^- = B_n \times T = B(x_0, r_n) \times (t', t_0),$$

and

$$k_n^- = \text{ess inf}_F u + \frac{\text{ess osc}_F u}{2^{\lambda+s+1}} + \frac{\text{ess osc}_F u}{2^{\lambda+s+n+1}},$$

$$A_n^- = \{(x, t) \in Q_n^- : (u(x, t) - k_n^-)_- > 0\},$$

where $s \geq 1$. Thus for every $n \geq 0$, we have

$$(u - k_n^-)_- \leq \varepsilon \text{ess osc}_F u, \quad \text{where } \varepsilon = \frac{1}{2^{\lambda+s}},$$

and for every $n = 0, 1, \dots$

$$(u - k_n^-)_-(x, t') = 0 \text{ for } \mu \text{ a.e. } x \in B(x_0, r_n).$$

By Lemma 4.4 we can now choose s so large that the corresponding initial condition (3.5)

$$\frac{\nu(A_0^-)}{\nu(Q_0^-)} \leq ((C_0(2 + 2^{(\lambda+1)(p-2)})^{2-p/\kappa})^{-1/(1-p/\kappa)} (4^{p2(2-p/\kappa)})^{1-(1-p/\kappa)^2})$$

is satisfied. By Lemma 3.4, we can then conclude that $\nu(A_{2n}^-)/\nu(Q_{2n}^-) \rightarrow 0$ as $n \rightarrow \infty$, which implies that

$$u \geq \text{ess inf}_F u + \frac{\text{ess osc}_F u}{2^{s+1}} \quad \text{a.e. in } B(x_0, r/8) \times (t', t_0),$$

for some s which depends only on the data and on λ . Since $t' \leq t_0 - \theta_-(r/4)^p$, the proof is complete. \square

Combining the above results, we obtain that the first alternative implies a reduction of oscillation in a subcylinder with upper time level at t_0 .

Corollary 4.7. *Assume that the first alternative holds. Then there exists a constant $\sigma_0 \in (0, 1)$ which depends only on the data and on λ , such that*

$$\operatorname{ess\,osc}_{B(x_0, \frac{r}{8}) \times (t_0 - \theta_-(\frac{r}{8})^p, t_0)} u \leq \sigma_0 \operatorname{ess\,osc}_F u.$$

Proof. As a consequence of Lemma 4.6, we know that there exists an $s \in \mathbb{N}$, which depends only on the data and on λ , such that

$$\begin{aligned} \operatorname{ess\,osc}_{B(x_0, \frac{r}{8}) \times (t_0 - \theta_-(\frac{r}{8})^p, t_0)} u &\leq \operatorname{ess\,sup}_F u - \operatorname{ess\,inf}_F u - \frac{\operatorname{ess\,osc}_F u}{2^{s+1}} \\ &= \left(1 - \frac{1}{2^{s+1}}\right) \operatorname{ess\,osc}_F u. \end{aligned}$$

□

This finishes the first alternative.

5. THE SECOND ALTERNATIVE

Suppose that for every $t^* \in (t_0 - \theta_+ r^p, t_0)$ we have

$$(5.1) \quad \begin{aligned} &\nu(\{B(x_0, r) \times (t^* - \theta_- r^p, t^*) : u(x, t) > k_0^-\}) \\ &\leq (1 - \alpha_0) \nu(B(x_0, r) \times (t^* - \theta_- r^p, t^*)), \end{aligned}$$

where α_0 is as in the first alternative and k_0^- is as in (4.1). This assumption is called *the second alternative*. Note that the second alternative is exactly the complement of the first alternative.

This alternative is also based on Lemma 3.1, but now we set

$$(5.2) \quad k_n^+ = \operatorname{ess\,sup}_F u - \frac{\operatorname{ess\,osc}_F u}{2^{\lambda+1}} - \frac{\operatorname{ess\,osc}_F u}{2^{\lambda+1+n}}$$

and will assume λ to be large enough so that we can force $\nu(A_0^+)$ to be small compared to $\nu(Q_0^+)$. We start with the following lemma.

Lemma 5.3. *Let u belong to the parabolic De Giorgi class and let the second alternative be in force. Let*

$$l = \operatorname{ess\,sup}_F u - \frac{\operatorname{ess\,osc}_F u}{2^{2s}}.$$

Then there exists a positive integer s which depends only on the data such that for almost every $t \in (t_0 - \theta_+ r^p, t_0)$

$$\mu(\{x \in B(x_0, r) : u(x, t) > l\}) \leq \frac{1 - \frac{3\alpha_0}{4}}{1 - \frac{\alpha_0}{2}} \mu(B(x_0, r)).$$

Proof. Let $t^* \in (t_0 - \theta_+ r^p, t_0)$. The second alternative implies that there exists a time level $t' \in (t^* - \theta_- r^p, t^* - \frac{\alpha_0}{2} \theta_- r^p)$ for which

$$(5.4) \quad \mu(\{x \in B(x_0, r) : u(x, t') > k_0^-\}) \leq \frac{1 - \alpha_0}{1 - \frac{\alpha_0}{2}} \mu(B(x_0, r)).$$

Indeed, if this was not the case, we would have

$$\begin{aligned} & \nu(\{(x, t) \in B(x_0, r) \times (t^* - \theta_- r^p, t^*) : u(x, t) > k_0^-\}) \\ & \geq \int_{t^* - \theta_- r^p}^{t^* - \frac{\alpha_0}{2} \theta_- r^p} \mu(\{x \in B(x_0, r) : u(x, t) > k_0^-\}) dt \\ & > (1 - \alpha_0) \nu(B(x_0, r) \times (t^* - \theta_- r^p, t^*)), \end{aligned}$$

which contradicts (5.1). Choose such a t' . Let s be a positive integer. We substitute

$$k = \operatorname{ess\,sup}_F u - \frac{\operatorname{ess\,osc}_F u}{2^s}$$

in (2.7) to obtain

$$\begin{aligned} & (l - k)_+^2 \mu(\{x \in B(x_0, (1 - \delta)r) : u(x, t) > l\}) \\ & \leq \operatorname{ess\,sup}_{t' < t < t^*} \int_{B(x_0, (1 - \delta)r)} (u(x, t) - k)_+^2 d\mu \\ & \leq C \int_{B(x_0, r)} (u(x, t') - k)_+^2 d\mu + \frac{C}{(\delta r)^p} \int_{B(x_0, r) \times (t^* - \theta_- r^p, t^*)} (u - k)_+^p d\nu \\ & \leq C \left(\frac{\operatorname{ess\,osc}_F u}{2^s} \right)^2 \mu(\{x \in B(x_0, r) : u(x, t') > k\}) \\ & \quad + \frac{C}{(\delta r)^p} \left(\frac{\operatorname{ess\,osc}_F u}{2^s} \right)^p \nu(B(x_0, r) \times (t^* - \theta_- r^p, t^*)) \end{aligned}$$

for almost every $t \in (t', t^*)$ and any $0 < \delta < 1$. With (5.4) and the definition of θ_- , this gives

$$\begin{aligned} & \mu(\{x \in B(x_0, (1 - \delta)r) : u(x, t) > l\}) \\ & \leq \left(\frac{\operatorname{ess\,osc}_F u}{(l - k)_+ 2^s} \right)^2 \frac{1 - \alpha_0}{1 - \frac{\alpha_0}{2}} \mu(B(x_0, r)) \\ & \quad + \frac{C}{(\delta r)^p} \left(\frac{\operatorname{ess\,osc}_F u}{(l - k)_+ 2^s} \right)^2 \left(\frac{1}{2^s} \right)^{p-2} r^p \mu(B(x_0, r)) \\ & \leq \left(1 - \frac{1}{2^s} \right)^2 \left(\frac{1 - \alpha_0}{1 - \frac{\alpha_0}{2}} + \frac{C}{\delta^p} \left(\frac{1}{2^s} \right)^{p-2} \right) \mu(B(x_0, r)), \end{aligned}$$

for almost every $t \in (t', t^*)$ and any $0 < \delta < 1$. By the α -annular decay property (2.1), we have

$$\begin{aligned} & \mu(\{x \in B(x_0, r) : u(x, t) > l\}) \\ & \leq \mu(B(x_0, r) \setminus B(x_0, (1 - \delta)r)) + \mu(\{x \in B(x_0, (1 - \delta)r) : u(x, t) > l\}) \\ & \leq C\delta^\alpha \mu(B(x_0, r)) + \mu(\{x \in B(x_0, (1 - \delta)r) : u(x, t) > l\}). \end{aligned}$$

Hence, by first choosing δ small enough and after this choosing s large enough, we obtain that for almost every $t \in (t^* - \frac{\alpha_0}{2} \theta_- r^p, t^*)$

$$(5.5) \quad \mu(\{x \in B(x_0, r) : u(x, t) > l\}) \leq \frac{1 - \frac{3\alpha_0}{4}}{1 - \frac{\alpha_0}{2}} \mu(B(x_0, r)).$$

The choices of δ and s depend only on the data. Finally, since the same choice of s is valid for every $t^* \in (t_0 - \theta_+ r^p, t_0)$, we conclude that (5.5) holds for almost every $t \in (t_0 - \theta_+ r^p, t_0)$. \square

Now we are ready to prove the final lemma which together with Lemma 3.4 gives the reduction of oscillation in case of the second alternative. For a constant k define

$$E_k^\tau = \{(x, t) \in B(x_0, \tau r) \times (t_0 - \theta_+ r^p, t_0) : u(x, t) > k\},$$

where $1 \leq \tau \leq 2$. For $\tau = 1$ we denote $E_k^1 = E_k$.

Lemma 5.6. *For every $0 < \alpha < 1$ there exists a positive constant λ , which depends only on the data and on α , such that*

$$\frac{\nu(E_{k_0^+})}{\nu(B(x_0, r) \times (t_0 - \theta_+ r^p, t_0))} \leq \alpha,$$

where as in (5.2),

$$k_0^+ = \operatorname{ess\,sup}_F u - \frac{\operatorname{ess\,osc}_F u}{2\lambda}.$$

Proof. Fix $0 < \alpha < 1$. Define

$$E_h^\tau(t) = \{x \in B(x_0, \tau r) : u(x, t) > h\},$$

and for constants h, k such that $h > k > k_0^+$, let

$$v = \begin{cases} h - k, & u \geq h, \\ u - k, & k < u < h, \\ 0, & u \leq k. \end{cases}$$

By the previous lemma we can choose λ big enough so that for almost every $t \in (t_0 - \theta_+ r^p, t_0)$ we have

$$\begin{aligned} \mu(\{x \in B(x_0, r) : v(x, t) = 0\}) &= \mu(\{x \in B(x_0, r) : u(x, t) \leq k\}) \\ &\geq \mu(\{x \in B(x_0, r) : u(x, t) < k_0^+\}) \\ &\geq \frac{\alpha_0}{4} \mu(B(x_0, r)). \end{aligned}$$

Thus for almost every $t \in (t_0 - \theta_+ r^p, t_0)$

$$v_{B(x_0, r)}(t) = \int_{B(x_0, r)} v(x, t) d\mu \leq \left(1 - \frac{\alpha_0}{4}\right) (h - k)$$

and consequently

$$h - k - v_{B(x_0, r)}(t) \geq \frac{\alpha_0}{4} (h - k).$$

Using the weak (q, q) -Poincaré inequality for some $q < p$, see Remark 2.5, gives

$$\begin{aligned} (h - k)^q \mu(E_h(t)) &\leq \left(\frac{4}{\alpha_0}\right)^q \int_{B(x_0, r)} |v(x, t) - v_{B(x_0, r)}(t)|^q d\mu \\ &\leq Cr^q \int_{B(x_0, \tau r) \times \{t\}} g_v^q d\mu = Cr^q \int_{E_k^\tau(t) \setminus E_h^\tau(t)} g_u^q d\mu \end{aligned}$$

for almost every $t \in (t_0 - \theta_+ r^p, t_0)$. Next we integrate the above inequality over time to get

$$(h - k)^q \nu(E_h) \leq Cr^q \int_{E_k^\tau \setminus E_h^\tau} g_u^q d\nu.$$

Now Hölder's inequality gives

$$\begin{aligned} (h - k)^q \nu(E_h) &\leq Cr^q \left(\int_{E_k^\tau \setminus E_h^\tau} g_u^p d\nu \right)^{q/p} \nu(E_k^\tau \setminus E_h^\tau)^{1-q/p} \\ &\leq Cr^q \left(\int_{B(x_0, \tau r) \times (t_0 - \theta_+ r^p, t_0)} g_{(u-k)_+}^p d\nu \right)^{q/p} \nu(E_k^\tau \setminus E_h^\tau)^{1-q/p}. \end{aligned}$$

Choose

$$k(s) = \operatorname{ess\,sup}_F u - \frac{\operatorname{ess\,osc}_F u}{2^s}.$$

By (2.7) and since $\mu(B(x_0, 2r))\theta_+ r^p = \nu(B(x_0, 2r) \times (t_0 - \theta_+ r^p, t_0))$, we obtain for every $s \leq \lambda$

$$\begin{aligned} (5.7) \quad &\int_{B(x_0, \tau r) \times (t_0 - \theta_+ r^p, t_0)} g_{(u-k)_+}^p d\nu \leq \int_{B(x_0, 2\tau r)} (u(x, t_0 - \theta_+ r^p) - k)_+^2 d\mu \\ &+ \frac{C}{(\tau r)^p} \int_{B(x_0, 2\tau r) \times (t_0 - \theta_+ r^p, t_0)} (u - k)_+^p d\nu \\ &\leq \frac{C}{r^p} \left(\frac{\operatorname{ess\,osc}_F u}{2^s} \right)^p \nu(B(x_0, r) \times (t_0 - \theta_+ r^p, t_0)). \end{aligned}$$

In the last step we also used the doubling property of μ . Choosing now

$$h(s) = \operatorname{ess\,sup}_F u - \frac{\operatorname{ess\,osc}_F u}{2^{s+1}},$$

yields

$$\nu(E_{h(s)}) \leq C \nu(B(x_0, r) \times (t_0 - \theta_+ r^p, t_0))^{q/p} \nu(E_{k(s)}^\tau \setminus E_{h(s)}^\tau)^{1-q/p}.$$

Finally, summing this over $s = 1, \dots, \lambda - 1$ and then using the doubling condition to replace τr by r gives

$$\begin{aligned} (\lambda - 1) \nu(E_{k_0^+})^{p/(p-q)} &\leq C \nu(B(x_0, r) \times (t_0 - \theta_+ r^p, t_0))^{q/(p-q)} \\ &\quad \cdot \nu(B(x_0, r) \times (t_0 - \theta_+ r^p, t_0)) \end{aligned}$$

and hence

$$\nu(E_{k_0^+}) \leq \frac{C}{(\lambda - 1)^{(p-q)/p}} \nu(B(x_0, r) \times (t_0 - \theta_+ r^p, t_0)).$$

Choosing λ large enough finishes the proof. \square

Now we are in the position to prove the reduction of oscillation in the case of the second alternative, and then complete the proof of the Hölder continuity of u .

Lemma 5.8. *Let u belong to the parabolic De Giorgi class and let the second alternative be in force. Then*

$$u \geq \operatorname{ess\,sup}_F u + \frac{\operatorname{ess\,osc}_F u}{2^{\lambda+1}} \quad \text{a.e. in } B(x_0, r/2) \times (t_0 - \theta_+ (r/2)^p, t_0),$$

where λ depends only on the data.

Proof. Set

$$r_n = \frac{r}{2} + \frac{r}{2^{n+1}}, \quad r_0 = r,$$

$$Q_n^+ = B_n \times T_n = B(x_0, r_n) \times (t_0 - \theta_+ r_n^p, t_0),$$

and

$$k_n^+ = \operatorname{ess\,sup}_F u - \frac{\operatorname{ess\,osc}_F u}{2^{\lambda+1}} - \frac{\operatorname{ess\,osc}_F u}{2^{\lambda+n+1}},$$

$$A_n^+ = \{ (x, t) \in Q_n^+ : u(x, t) > k_n^+ \}.$$

Hence the corresponding condition Then for every $n \geq 0$, we have

$$(u - k_n^-)_- \leq \varepsilon \operatorname{ess\,osc}_F u, \quad |k_{n+1}^\pm - k_n^\pm| \geq \varepsilon \frac{\operatorname{ess\,osc}_F u}{2^{n+2}}, \quad \text{where } \varepsilon = \frac{1}{2^\lambda}.$$

Hence, the initial condition (3.5) corresponding to Lemma 3.1 takes on the form

$$\frac{\nu(A_0^+)}{\nu(Q_0^+)} \leq (C_0(3 + 1^{p-2} + 1^{2-p})^{2-p/\kappa})^{-1/(1-p/\kappa)} (4^{p2(2-p/\kappa)})^{1-(1-p/\kappa)^2}.$$

Since the right hand side depends only on the data, by Lemma 5.6 we see that there exists a λ , which depends only on the data, for which the above condition is satisfied. Assume λ is such. Then as $n \rightarrow 0$ we have $\nu(A_{2n}^+)/\nu(Q_{2n}^+) \rightarrow 0$, which implies the statement of the lemma. \square

Corollary 5.9. *Suppose that the second alternative holds. Then there exists a $\sigma_1 \in (0, 1)$, which depends only on the data, such that*

$$\operatorname{ess\,osc}_{B(x_0, \frac{r}{2}) \times (t_0 - \theta_+ (\frac{r}{2})^p, t_0)} u \leq \sigma_1 \operatorname{ess\,osc}_F u.$$

Proof. By the previous lemma, we have

$$\operatorname{ess\,sup}_{B(x_0, \frac{r}{2}) \times (t_0 - \theta_+ (\frac{r}{2})^p, t_0)} u \leq \operatorname{ess\,sup}_F u - \frac{\operatorname{ess\,osc}_F u}{2^{\lambda+1}}$$

for some $\lambda > 1$ which depends only upon the data. This implies the statement of the lemma with

$$\sigma_1 = 1 - \frac{1}{2^{\lambda+1}}.$$

\square

Proof of Theorem 2.8. Since either the first alternative or the second alternative is in force, by Corollary 4.7 and Corollary 5.9 we know that

$$\operatorname{ess\,osc}_{B(x_0, \frac{r}{8}) \times (t_0 - \theta_- (\frac{r}{8})^p, t_0)} u \leq \sigma \operatorname{ess\,osc}_F u,$$

where $\sigma = \max\{\sigma_0, \sigma_1\} < 1$ and θ_- depend only on the data. The local Hölder continuity of u now follows from this reduction of oscillation, by a standard recursive argument presented for example in [Urb08], p.44.

6. REGULARITY OF PARABOLIC QUASIMINIMIZERS

In this section we show that parabolic quasiminimizers belong to the parabolic De Giorgi class. By the previous sections, bounded parabolic quasiminimizers are thus locally Hölder continuous.

6.1. Parabolic quasiminimizers.

Definition 6.1. A function $u \in L^p_{\text{loc}}(0, T; N^{1,p}_{\text{loc}}(\Omega))$ is a parabolic K -quasiminimizer if there exists a constant $K \geq 1$ and a Carathéodory function $G(x, t, u, g)$, satisfying the growth condition

$$(6.2) \quad C_1|q|^p \leq G(x, t, v, q) \leq C_2|q|^p$$

for some positive constants C_1 and C_2 , such that

$$(6.3) \quad - \int_F u \frac{\partial \phi}{\partial t} d\nu + E(u; F) \leq K E(u - \phi; F)$$

for every open $F \subset \subset \Omega_T$ and $\phi \in C^\infty(0, T; N^{1,p}(\Omega))$ such that $\{\phi \neq 0\} \subset F$. Here

$$E(u; F) = \int_F G(x, t, u, g_u) d\nu.$$

By (6.2) and (6.3) we have that if u is a K -quasiminimizer, then there exists positive constants $0 < C_1 < C_2$ such that

$$(6.4) \quad - \int_F u \frac{\partial \phi}{\partial t} d\nu + C_1 \int_F g_u^p d\nu \leq K C_2 \int_F g_{u-\phi}^p d\nu,$$

for every open $F \subset \subset \Omega_T$ and $\phi \in C^\infty(0, T; N^{1,p}(\Omega))$ such that $\{\phi \neq 0\} \subset F$. This implies the following.

Lemma 6.5. *Let $u \in L^p_{\text{loc}}(0, T; N^{1,p}_{\text{loc}}(\Omega))$ be a parabolic K -quasiminimizer. Then there exist $0 < C_1 < C_2$ such that*

$$- \int_{\{\phi \neq 0\}} u \frac{\partial \phi}{\partial t} d\nu + C_1 \int_{\{\phi \neq 0\}} g_u^p d\nu \leq K C_2 \int_{\{\phi \neq 0\}} g_{u-\phi}^p d\nu$$

for every $\phi \in C^\infty(0, T; N^{1,p}(\Omega))$ such that $\text{supp } \phi \subset \subset \Omega$.

Proof. Let $\varepsilon > 0$ and $\phi \in C^\infty(0, T; N^{1,p}(\Omega))$ be such that $\text{supp } \phi \subset \subset \Omega$. Since $\{\phi \neq 0\}$ is ν -measurable and compactly contained in Ω_T , and since $g_u \in L^p(\Omega_T)$, there exists an open set $F \subset \subset \Omega_T$ such that

$$\int_{F \setminus \{\phi \neq 0\}} g_u^p d\nu \leq \frac{\varepsilon}{K C_2}.$$

Also, since ϕ is continuous with respect to time we have $\nu(\{\phi = 0, \partial \phi / \partial t \neq 0\}) = 0$, and so by (6.4) we can write

$$\begin{aligned} - \int_{\{\phi \neq 0\}} u \frac{\partial \phi}{\partial t} d\nu + C_1 \int_{\{\phi \neq 0\}} g_u^p d\nu &\leq - \int_F u \frac{\partial \phi}{\partial t} d\nu + C_1 \int_F g_u^p d\nu \\ K C_2 \int_F g_{u-\phi}^p d\nu &\leq K C_2 \int_{\{\phi \neq 0\}} g_{u-\phi}^p d\nu + \varepsilon. \end{aligned}$$

This holds for every $\varepsilon > 0$, which completes the proof. \square

Our aim in what follows is to prove that a parabolic quasiminimizer u belongs to the parabolic De Giorgi Class, i.e. fulfills the estimate of Definition 2.6.

A fundamental part of the proof is to use partial integration on u with respect to the time variable. However, the time regularity of the function

$u \in L_{\text{loc}}^p(0, T; N_{\text{loc}}^{1,p}(\Omega))$ is a priori not sufficient for this. Therefore we first establish suitable estimates for u_ε , which denotes the standard time mollification of u . Having done this we then pass to the limit as $\varepsilon \rightarrow 0$, and obtain the desired results for u .

In the Euclidean setting this argument works, since the theory of mollifiers together with the linearity of taking a gradient guarantees that $u_\varepsilon \rightarrow u$ and $\nabla u_\varepsilon \rightarrow \nabla u$ in $L_{\text{loc}}^p(\Omega_T)$ as $\varepsilon \rightarrow 0$. Unfortunately however, the presence of upper gradients in place of usual gradients causes complication to the argument. Since the operation of taking an upper gradient is not linear, it turns out to be problematic to show that $g_{u-u_\varepsilon} \rightarrow 0$ in $L_{\text{loc}}^p(\Omega_T)$ as $\varepsilon \rightarrow 0$. It would be interesting to know whether or not it is possible to show this using only the theory of upper gradients.

We circumvent this question by using the known comparability between upper gradients and so called Cheeger derivatives. As will be seen in the following, the Cheeger derivative has the property of being a linear operation.

6.2. The Cheeger derivative. The following theorem, which yields in a local sense the notion of partial derivatives in metric space, is by Cheeger [Che99]. For a concise source of tools given by the theory of Cheeger derivatives we refer to [BBS03] and the references therein.

Theorem 6.6. *Let X be a metric measure space equipped with a positive doubling Borel regular measure μ . Assume X admits a weak $(1, p)$ -Poincaré inequality for some $1 < p < \infty$.*

Then there exists a countable collection (U_α, X^α) of measurable sets U_α and Lipschitz functions $X^\alpha = (X_1^\alpha, \dots, X_{k(\alpha)}^\alpha) : X \rightarrow \mathbb{R}^{k(\alpha)}$ such that $\mu(X \setminus \bigcup_\alpha U_\alpha) = 0$ and for all α , the following hold:

The functions $X_1^\alpha, \dots, X_{k(\alpha)}^\alpha$ are linearly independent on U_α and $1 \leq k(\alpha) \leq N$, where N is a constant depending only on the doubling constant of μ and the constant from the Poincaré inequality. If $f : X \rightarrow \mathbb{R}$ is Lipschitz, then there exist unique measurable bounded vector valued functions $d^\alpha f : U_\alpha \rightarrow \mathbb{R}^{k(\alpha)}$ such that for μ -a.e. $x_0 \in U_\alpha$,

$$\lim_{r \rightarrow 0^+} \sup_{x \in B(x_0, r)} \frac{|f(x) - f(x_0) - \langle d^\alpha f(x_0), X^\alpha(x) - X^\alpha(x_0) \rangle|}{r} = 0.$$

A non negative function $|\cdot|_{1, x_0}$ is introduced [Che99], p.460, on $d_\alpha f(x_0)$ such that

$$|d^\alpha f(x_0)|_{1, x_0} = g_f(x_0),$$

where g_f is the minimal p -weak upper gradient of f . Furthermore, it is shown that one can find an inner product norm

$|\cdot|_{x_0} : \mathbb{R}^{k(\alpha)} \rightarrow [0, \infty)$ which is C -quasi-isometric to $|\cdot|_{1, x_0}$, where the constant C depends only on $k(\alpha)$.

We may assume that the sets U_α are pairwise disjoint. For each α , extend $d^\alpha f$ to be zero in the set $X \setminus U_\alpha$, and define

$$Df : X \rightarrow \mathbb{R}^{k(\alpha)},$$

$$Df = \sum_{\alpha} d^\alpha f.$$

The above imply that the differential mapping $D : f \mapsto Df$ is linear, and that there is a constant $C > 0$, which depends only on N , such that for all Lipschitz functions f and μ -a.e. $x \in X$

$$(6.7) \quad \frac{1}{C}|Df(x)| \leq g_f(x) \leq C|Df(x)|,$$

where by $|Df(x)|$ we mean $|d^\alpha f(x)|_x$, whenever $x \in U_\alpha$.

From [Sha00] it is known that the Newtonian space $N^{1,p}(X)$ is the closure in the $N^{1,p}$ -norm, of the collection of Lipschitz functions on X with finite $N^{1,p}$ -norm. By [FHP99] we know that there exists a unique gradient Du which satisfies (6.7) for every $u \in N^{1,p}(X)$. Also, if $\{u_j\}_{j=1}^\infty$ is a sequence in $N^{1,p}(X)$, then $u_j \rightarrow u$ in $N^{1,p}(X)$ if and only if $u_j \rightarrow u$ and $Du_j \rightarrow Du$ in $L^p(X, \mu; \mathbb{R}^N)$, as $j \rightarrow \infty$.

Analogously to what was done with upper gradients, we define the parabolic Cheeger derivative of a time dependent function by taking the Cheeger derivative with respect to the variable x , at time level t .

Next we prove the steps which will be used to overcome the complications in the mollification argument, caused by the non linearity of upper gradients. Here and in what follows we denote by u_h the time mollification of a function u , i.e.

$$u_h(x, t) = \int_{-h}^h \eta_h(s) u(x, t-s) ds,$$

where $\eta_h(s) = h^{-1}\eta(s/h)$ denotes a standard mollifier.

Lemma 6.8. *Assume $u \in L^p_{loc}(0, T; N^{1,p}_{loc}(\Omega))$. Then as $h \rightarrow 0$, it holds $g_{u-u_h} \rightarrow 0$ in $L^p_{loc}(\Omega_T)$, and also pointwise ν -almost everywhere in Ω_T . Moreover, as $s \rightarrow 0$, we have $g_{u(\cdot, -s)-u} \rightarrow 0$ in $L^p_{loc}(\Omega_T)$.*

Proof. Let $h > 0$. For ν -almost every $x, y \in \Omega$, $t \in [t_1 + h, t_2 - h]$ and every compact rectifiable path γ from x to y , we have

$$\begin{aligned} |u_h(x, t) - u_h(y, t)| &\leq \int_{t_1}^{t_2} \left(\int_\gamma (g_u)(z, s) dz \right) \eta_h(t-s) ds \\ &= \int_\gamma (g_u)_h(z, t) dz. \end{aligned}$$

Hence $(g_u)_h$ is a p -weak upper gradient of u_h . The definition of the minimal p -weak upper gradient now implies that for ν -almost every $(x, t) \in \Omega \times [t_1 + h, t_2 - h]$

$$(6.9) \quad g_{u_h}(x, t) \leq (g_u)_h(x, t).$$

We now show that for ν -every $(x, t) \in \Omega \times (h, T-h)$ we have $Du_h(x, t) = (Du)_h(x, t)$. Assume first that $u \in \text{Lip}(\Omega \times (0, T))$. Assume a point $(x_0, t) \in \Omega \times (0, T)$ where with respect to the spatial variable, the Cheeger derivative of u exists. Let $r > 0$ be such that $B(x_0, r) \subset \Omega$. Then, since u and X^α are Lipschitz-continuous with modulus $C_{\text{Lip}(u)}$ and $C_{\text{Lip}(X^\alpha)}$ respectively, we

may write for any $x \in B(x_0, r)$

$$\begin{aligned} & r^{-1} |u_h(x, t) - u_h(x_0, t) - \langle (d^\alpha u)_h(x_0, t), X^\alpha(x) - X^\alpha(x_0) \rangle| \\ & \leq \int_{-h}^h \eta_h(t-s) r^{-1} |u(x, s) - u(x_0, s) - \langle (d^\alpha u)(x_0, s), X^\alpha(x) - X^\alpha(x_0) \rangle| ds \\ & \leq C \cdot C_{\text{Lip}(u)} + C \cdot C_{\text{Lip}(u)} \cdot C_{\text{Lip}(X^\alpha)}. \end{aligned}$$

Therefore by the Lebesgue theorem of dominated convergence we have

$$\begin{aligned} & \lim_{r \rightarrow 0^+} \sup_{x \in B(x_0, r)} r^{-1} |u_h(x, t) - u_h(x_0, t) - \langle (d^\alpha u)_h(x_0, t), X^\alpha(x) - X^\alpha(x_0) \rangle| \\ & \leq \int_{-h}^h \eta_h(t-s) \\ & \cdot \lim_{r \rightarrow 0^+} \sup_{x \in B(x_0, r)} r^{-1} |u(x, s) - u(x_0, s) - \langle (d^\alpha u)(x_0, s), X^\alpha(x) - X^\alpha(x_0) \rangle| ds \\ & = 0. \end{aligned}$$

By the uniqueness of the Cheeger derivative, and by the definition of Du the above implies that $Du_h(x_0, t) = (Du)_h(x_0, t)$. Assume then that $u \in L_{\text{loc}}^p(0, T; N_{\text{loc}}^{1,p}(\Omega))$, not necessarily Lipschitz, and let F be a compact subset of Ω_T . Let $\{u_j\} \subset \text{Lip}(\Omega_T)$ be a sequence such that $u_j \rightarrow u$ in $L^p(0, T; N^{1,p}(\Omega))$. Then, since u_j is Lipschitz, by inequality (6.7) and by (6.9)

$$\begin{aligned} & \|Du_h - (Du)_h\|_{L^p(F)} \leq \|D(u - u_j)_h\|_{L^p(F)} + \|D(u_j)_h - (Du)_h\|_{L^p(F)} \\ & \leq C \|g_{(u-u_j)_h}\|_{L^p(F)} + \|(Du_j)_h - (Du)_h\|_{L^p(F)} \\ & \leq C \|(g_{u-u_j})_h\|_{L^p(F)} + \|(Du_j - Du)_h\|_{L^p(F)} \\ & \leq C \|g_{u-u_j}\|_{L^p(F)} + C \|Du_j - Du\|_{L^p(F)}. \end{aligned}$$

Since the last expression tends to zero as $j \rightarrow \infty$, we can conclude that $Du_h(x, t) = (Du)_h(x, t)$ for ν -almost every $(x, t) \in \Omega \times (h, T-h)$.

By inequality (6.7), by the linearity of the Cheeger derivation and since $Du_h = (Du)_h$, we can write for ν -almost every $(x, t) \in \Omega \times (h, T-h)$

$$(6.10) \quad g_{u-u_h} \leq C |Du - (Du)_h|.$$

Since $g_u \in L_{\text{loc}}^p(\Omega_T)$, by (6.7), also $Du \in L_{\text{loc}}^p(\Omega_T)$. This means, by the theory of mollifiers, that as $h \rightarrow 0$, on the right side of (6.10) we have convergence to zero in $L_{\text{loc}}^p(\Omega_T)$ and also pointwise ν -almost everywhere in Ω_T . Lastly, for an $s > 0$ small enough, by inequality (6.7) we have

$$\int_F g_{u(\cdot, \cdot-s)-u}^p d\nu \leq C \int_F |Du(\cdot, \cdot-s) - Du|^p d\nu.$$

Again by (6.7), we know that $Du \in L_{\text{loc}}^p(\Omega_T)$. Since F is compact, the Lemma now follows from the continuity of the translation operation for L^p functions. \square

Now we are set to prove that parabolic quasiminimizers belong to the De Giorgi class.

Theorem 6.11. *Let $u \in L_{\text{loc}}^p(0, T; N_{\text{loc}}^{1,p}(\Omega))$ be a parabolic K -quasiminimizer. Then u belongs to the parabolic De Giorgi class.*

Proof. Let u be a parabolic quasiminimizer. By making a change of variable, it is straightforward to check that if $u(x, t)$ fulfills (6.4), then for any small enough s also $u(x, t - s)$ fulfills (6.4). Assume a function $\phi \in L^p(0, T; N^{1,p}(\Omega))$ such that $\text{supp } \phi \subset \subset \Omega_T$. Then there exists an $h_0 > 0$ such that for every $0 < h < h_0$, $\phi_h \in C^\infty(0, T; N^{1,p}(\Omega))$ with $\text{supp } \phi_h \subset \subset \Omega_T$, and so is a permissible test function. By Lemma 6.5, we have

$$(6.12) \quad \begin{aligned} & - \int_{\{\phi_h \neq 0\}} u(x, t - s) \frac{\partial \phi_h}{\partial t} d\nu + C_1 \int_{\{\phi_h \neq 0\}} g_{u(\cdot, \cdot - s)}^p d\nu, \\ & \leq KC_2 \int_{\{\phi_h \neq 0\}} g_{u(\cdot, \cdot - s) - \phi_h}^p d\nu. \end{aligned}$$

We multiply both sides of (6.12) by a standard mollifier with respect to the variable s and which has support $(-\varepsilon, \varepsilon)$. Integrating the resulting expression in the variable s yields, after using Fubini's theorem,

$$\begin{aligned} & - \int_{\{\phi_h \neq 0\}} u_\varepsilon(x, t) \frac{\partial \phi_h}{\partial t} d\nu + C_1 \int_{\{\phi_h \neq 0\}} (g_{u(\cdot, \cdot - s)}^p)_\varepsilon d\nu \\ & \leq KC_2 \int_{\{\phi_h \neq 0\}} (g_{u(\cdot, \cdot - s) - \phi_h}^p)_\varepsilon d\nu. \end{aligned}$$

Conducting partial integration on the time derivative term and using the triangle inequality for upper gradients on the right side yields

$$(6.13) \quad \begin{aligned} & \int_{\{\phi_h \neq 0\}} \frac{\partial u_\varepsilon}{\partial t} \phi_h d\nu + C_1 \int_{\{\phi_h \neq 0\}} (g_{u(\cdot, \cdot - s)}^p)_\varepsilon d\nu \\ & \leq KC \int_{\{\phi_h \neq 0\}} (g_{u(\cdot, \cdot - s) - \phi}^p)_\varepsilon d\nu + KC \int_{\{\phi_h \neq 0\}} g_{\phi - \phi_h}^p d\nu. \end{aligned}$$

As $h \rightarrow 0$, from the theory of mollifiers, it follows that $\{\phi_h \neq 0\}$ converges to $\{\phi \neq 0\}$ in ν -measure as $h \rightarrow 0$. By Lemma 6.8 the last term on the right hand side of (6.13) converges to zero as $h \rightarrow 0$. Hence after taking the limit $h \rightarrow 0$ we have

$$(6.14) \quad \begin{aligned} & \int_{\{\phi \neq 0\}} \frac{\partial u_\varepsilon}{\partial t} \phi d\nu + C_1 \int_{\{\phi \neq 0\}} (g_{u(\cdot, \cdot - s)}^p)_\varepsilon d\nu \\ & \leq KC \int_{\{\phi \neq 0\}} (g_{u(\cdot, \cdot - s) - \phi}^p)_\varepsilon d\nu. \end{aligned}$$

Assume now $x \in \Omega$ and $r_1 < r_2$ and $\tau_2 < \tau_1 < \tau_0$ to be such that $B(x, r_2) \times (\tau_2, \tau_0) \subset \Omega_T$. Let $\chi_{[\tau_2, t]}$ denote the characteristic function of the time interval $[\tau_2, t]$. Let $\varphi \in \text{Lip}(\Omega_T)$ be such that $0 \leq \varphi \leq 1$, $\varphi = 1$ in $B(x, r_1) \times (\tau_1, \tau_0)$, that $\varphi(x, t) = 0$ whenever $t \leq \tau_2$ or $x \notin B(x, r_2)$, and

$$g_\varphi^p \leq \frac{1}{(r_2 - r_1)^p} \quad \text{and} \quad \left| \frac{\partial \varphi}{\partial t} \right| \leq \frac{1}{\tau_1 - \tau_2}.$$

Choose the test function

$$\phi = \pm \varphi(u_\varepsilon - k) \pm \chi_{[\tau_2, t]},$$

where t is arbitrarily fixed in (τ_2, τ_0) . We can now estimate on the right hand side of (6.14),

$$\begin{aligned} \int_{\{\phi \neq 0\}} (g_{u(\cdot, \cdot - s) \mp \varphi(u_\varepsilon - k) \pm \chi_{[\tau_2, t]}}^p)_\varepsilon d\nu &\leq \left(C \int_{\{\phi \neq 0\}} g_{u(\cdot, \cdot - s) - u}^p d\nu \right)_\varepsilon \\ &+ C \int_{\{\phi \neq 0\}} g_{u - u_\varepsilon}^p d\nu + C \int_{\{\phi \neq 0\}} g_{u_\varepsilon - \varphi(u_\varepsilon - k)}^p d\nu. \end{aligned}$$

By Lemma 6.8, the first and second terms on the right hand side of the above expression converge to zero as $\varepsilon \rightarrow 0$. For the third term we write

$$\begin{aligned} \int_{\{\phi \neq 0\}} g_{u_\varepsilon - \varphi(u_\varepsilon - k)}^p d\nu &= \int_{\{\phi \neq 0\}} g_{(1 - \varphi)(u_\varepsilon - k)}^p d\nu \leq C \int_{\{\phi \neq 0\}} (1 - \varphi)^p g_{u_\varepsilon - u}^p d\nu \\ &+ C \int_{\{\phi \neq 0\}} (1 - \varphi)^p g_u^p d\nu + C \int_{\{\phi \neq 0\}} (u_\varepsilon - k)_\pm^p g_\varphi^p d\nu \end{aligned}$$

Hence, as $\varepsilon \rightarrow 0$, after recalling the properties of φ , and the definition of ϕ , we obtain for the right hand side of (6.14),

$$\begin{aligned} \limsup_{\varepsilon \rightarrow 0} \int_{\{\phi \neq 0\}} (g_{u(\cdot, \cdot - s) - \phi}^p)_\varepsilon d\nu &\leq C \int_{\tau_2}^{\tau_1} \int_{B(x, r_2) \setminus B(x, r_1)} g_{(u - k)_\pm}^p d\nu \\ &+ \frac{C}{(r_2 - r_1)^p} \int_{\tau_2}^{\tau_0} \int_{B(x, r_2)} (u - k)_\pm^p d\nu. \end{aligned}$$

On the left hand side of (6.14), for the first term we have

$$\begin{aligned} &\pm \int_{\{\phi \neq 0\}} \frac{\partial u_\varepsilon}{\partial t} \varphi(u_\varepsilon - k) \chi_{[\tau_2, t]} d\nu = \frac{1}{2} \int_{\{\phi \neq 0\}} \frac{\partial}{\partial t} (\varphi(u_\varepsilon - k)_\pm^2) \chi_{[\tau_2, t]} d\nu \\ &\quad - 2 \int_{\{\phi \neq 0\}} \frac{\partial \varphi}{\partial t} (u_\varepsilon - k)_\pm^2 \chi_{[\tau_2, t]} d\nu \\ &\longrightarrow \frac{1}{2} \int_{B(x, r_2)} \varphi(u(x, t) - k)_\pm^2 d\mu - 2 \int_{\tau_2}^t \int_{B(x, r_2)} \frac{\partial \varphi}{\partial t} (u - k)_\pm^2 d\nu, \end{aligned}$$

as $\varepsilon \rightarrow 0$. For the second term on the left hand side of (6.14), we clearly have

$$\lim_{\varepsilon \rightarrow 0} \int_{\{\phi \neq 0\}} (g_{u(\cdot, \cdot - s)}^p)_\varepsilon d\nu \geq \int_{\tau_2}^t \int_{B(x, r_1)} g_{(u - k)_\pm}^p d\nu.$$

Collecting the results, since the constants of the obtained inequality are independent of t , we obtain the estimate

$$\begin{aligned} &\text{ess sup}_{\tau_1 < t < \tau_0} \int_{B(x, r_1)} (u(x, t) - k)_\pm^2 d\mu + \int_{\tau_2}^{\tau_0} \int_{B(x, r_1)} g_{(u - k)_\pm}^p d\nu \\ &\leq C \int_{\tau_2}^{\tau_0} \int_{B(x, r_2) \setminus B(x, r_1)} g_{(u - k)_\pm}^p d\nu + \frac{C}{(r_2 - r_1)^p} \int_{\tau_2}^{\tau_0} \int_{B(x, r_2)} (u - k)_\pm^p d\nu \\ &\quad + \frac{C}{\tau_1 - \tau_2} \int_{\tau_2}^{\tau_1} \int_{B(x, r_2)} (u - k)_\pm^2 d\nu. \end{aligned}$$

We then multiply the second term on the left hand side by the constant C , and sum the resulting term on both sides of the above inequality, to obtain

$$\begin{aligned}
 & \operatorname{ess\,sup}_{\tau_1 < t < \tau_0} \int_{B(x, r_1)} (u(x, t) - k)_\pm^2 d\mu + C_1 \int_{\tau_1}^{\tau_0} \int_{B(x, r_1)} g_{(u-k)_\pm}^p d\nu \\
 (6.15) \quad & \leq C_2 \int_{\tau_2}^{\tau_0} \int_{B(x, r_2)} g_{(u-k)_\pm}^p d\nu + \frac{C_2}{(r_2 - r_1)^p} \int_{\tau_2}^{\tau_0} \int_{B(x, r_2)} (u - k)_\pm^p d\nu \\
 & \quad + \frac{C_2}{\tau_1 - \tau_2} \int_{\tau_2}^{\tau_1} \int_{B(x, r_2)} (u - k)_\pm^2 d\nu,
 \end{aligned}$$

where $C_1 > C_2 > 0$. Set now

$$f(\rho, \tau) = \frac{1}{C_1} \operatorname{ess\,sup}_{\tau < t < \tau_0} \int_{B(x, \rho)} (u(x, t) - k)_\pm^2 d\mu + \int_{\tau}^{\tau_0} \int_{B(x, \rho)} g_{(u-k)_\pm}^p d\nu.$$

Then by (6.15), for every $0 < r_1 < r_2 < r$ and $0 < \tau_2 < \tau_1 < \tau_0$, we have

$$\begin{aligned}
 f(r_1, \tau_1) & \leq \frac{C_2}{C_1} f(r_2, \tau_2) \\
 & + \frac{C}{(r_2 - r_1)^p} \int_{\tau_2}^{\tau_0} \int_{B(x, r_2)} (u - k)_\pm^p d\nu + \frac{C}{(\tau_1 - \tau_2)} \int_{\tau_2}^{\tau_1} \int_{B(x, r_2)} (u - k)_\pm^2 d\nu.
 \end{aligned}$$

Using the following real analytic lemma now completes the proof. \square

Lemma 6.16. *Let $f(\rho, \tau)$ be a nonnegative and bounded function on $[0, r] \times [0, \tau_0]$. If for every $0 < r_1 < r_2 < r$ and $0 < \tau_2 < \tau_1 < \tau_0$, we have*

$$f(r_1, \tau_1) \leq \sigma_3 f(r_2, \tau_2) + A(r_2 - r_1)^{-\alpha} + B(\tau_1 - \tau_2)^{-\beta},$$

where $A, B, \alpha, \beta, \sigma_3$ are nonnegative constants and $0 \leq \sigma_3 < 1$, then for every $0 < r_1 < r_2 < r$ and $0 < \tau_2 < \tau_1 < \tau_0$

$$f(r_1, \tau_1) \leq CA(r_2 - r_1)^{-\alpha} + CB(\tau_1 - \tau_2)^{-\beta},$$

where the constant C depends only on α, β, σ_3 .

Proof. For proof use Lemma 2.1.4 from [WZYL01] for both variables, one at a time. \square

REFERENCES

- [BB11] Anders Björn and Jana Björn. *Nonlinear Potential Theory on Metric Spaces*, volume 17 of *EMS Tracts in Mathematics*. European Mathematical Society (EMS), Zürich, 2011.
- [BBS03] Anders Björn, Jana Björn, and Nageswari Shanmugalingam. The diriclet problem for p -harmonic functions on metric spaces. *J.Reine Angew. Math*, 556:173–203, 2003.
- [Buc99] Stephen M. Buckley. Is the maximal function of a Lipschitz function continuous? *Ann. Acad. Sci. Fenn. Math.*, 24(2):519–528, 1999.
- [Che99] Jeff Cheeger. Differentiability of lipschitz functions on metric measure spaces. *Geom. Funct. Anal.*, 9(3):428–517, 1999.
- [CS84a] Filippo Chiarenza and Raul Serapioni. Degenerate parabolic equations and Harnack inequality. *Ann. Mat. Pura Appl. (4)*, 137:139–162, 1984.
- [CS84b] Filippo M. Chiarenza and Raul P. Serapioni. A Harnack inequality for degenerate parabolic equations. *Comm. Partial Differential Equations*, 9(8):719–749, 1984.

- [CS85] Filippo Chiarenza and Raul Serapioni. A remark on a Harnack inequality for degenerate parabolic equations. *Rend. Sem. Mat. Univ. Padova*, 73:179–190, 1985.
- [DiB86] Emmanuele DiBenedetto. On the local behaviour of solutions of degenerate parabolic equations with measurable coefficients. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 13(3):487–535, 1986.
- [DiB88] Emmanuele DiBenedetto. Harnack estimates in certain function classes. *Atti seminario matematico e fisico univ. Modena*, 37:173–182, 1988.
- [DiB93] Emmanuele DiBenedetto. *Degenerate parabolic equations*. Universitext. Springer-Verlag, New York, 1993.
- [FHP99] B. Franchi, P. Hajlasz, and Koskela P. Definitions of sobolev classes in metric spaces. *Ann. Inst. Fourier(Grenoble)*, 49:1903–1924, 1999.
- [Gia93] Mariano Giaquinta. *Introduction to regularity theory for nonlinear elliptic systems*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 1993.
- [Giu] Enrico Giusti. *Direct methods in the calculus of variations*.
- [GV06] Ugo Gianazza and Vincenzo Vespri. Parabolic De Giorgi classes of order p and the Harnack inequality. *Calc. Var. Partial Differential Equations*, 26(3):379–399, 2006.
- [Hei01] Juha Heinonen. *Lectures on analysis on metric spaces*. Universitext. Springer-Verlag, New York, 2001.
- [HK95] Piotr Hajlasz and Pekka Koskela. Sobolev meets Poincaré. *C. R. Acad. Sci. Paris*, 320(1):1211–1215, 1995.
- [HK98] Juha Heinonen and Pekka Koskela. Quasiconformal maps in metric spaces with controlled geometry. *Acta Math.*, 181:1–61, 1998.
- [KLSU11] Tuomo Kuusi, Rojbin Laleoglu, Juhana Siljander, and José Miguel Urbano. Hölder continuity for trudinger’s equation in measure spaces. To appear in *Calc. Var. Partial Differential Equations*, 2011.
- [KM98] P. Koskela and P. Macmanus. Quasiconformal mappings and sobolev spaces. *Studia Math.*, 131:1–17, 1998.
- [KS01] Juha Kinnunen and Nageswari Shanmugalingam. Regularity of quasinimizers on metric spaces. *Manuscripta Math.*, 105(3):401–423, 2001.
- [KSU11] Tuomo Kuusi, Juhana Siljander, and José Miguel Urbano. Local hölder continuity for doubly nonlinear parabolic equations. To appear in *Indiana Univ. Math. J.*, 2011.
- [KZ08] Stephen Keith and Xiao Zhong. The Poincaré inequality is an open ended condition. *Ann. of Math. (2)*, 167(2):575–599, 2008.
- [LSU68] O.A. Ladyzhenskaja, V.A. Solonnikov, and N.N. Ural’ceva. Linear and quasilinear equations of parabolic type. *Amer. Math. Soc. Transl. Math. Mono.*, 23, 1968.
- [Sha00] Nageswari Shanmugalingam. Newtonian spaces: an extension of Sobolev spaces to metric measure spaces. *Rev. Mat. Iberoamericana*, 16(2):243–279, 2000.
- [Sha01] Nageswari Shanmugalingam. Harmonic functions on metric spaces. *Illinois J. Math.*, 45:1021–1050, 2001.
- [Urb08] José Miguel Urbano. *The method of intrinsic scaling*, volume 1930 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2008. A systematic approach to regularity for degenerate and singular PDEs.
- [Wie87] W. Wieser. Parabolic Q -minima and minimal solutions to variational flows. *Manuscripta Math.*, 59(1):63–107, 1987.
- [WZYL01] Z. Wu, J. Zhao, J. Yin, and H. Li. *Nonlinear Diffusion Equation*. World Scientific, 2001.
- [Zho93] Shulin Zhou. On the local behaviour of parabolic Q -minima. *J. Partial Differential Equations*, 6(3):255–272, 1993.
- [Zho94] Shulin Zhou. Parabolic Q -minima and their application. *J. Partial Differential Equations*, 7(4):289–322, 1994.

(M.M) AALTO UNIVERSITY, DEPARTMENT OF MATHEMATICS AND SYSTEMS ANALYSIS,
P.O. BOX 11100 FI-00076 AALTO, FINLAND

E-mail address: `mathiasmasson@hotmail.com`

(J.S.) UNIVERSITY OF HELSINKI, DEPARTMENT OF MATHEMATICS AND STATISTICS,
P.O. BOX 68, FI-00014 UNIVERSITY OF HELSINKI, FINLAND

E-mail address: `juhana.siljander@helsinki.fi`